ON JORDAN OPERATORS AND RIGIDITY OF LINEAR CONTROL SYSTEMS

by H. O. FATTORINI*
University of Buenos Aires

INTRODUCTION

Let $E$ be a vector space over a field $K$. A linear operator $A$ in $E$ will be called a Jordan operator if there exists a non-null polynomial $P(\lambda) = a_0 + a_1 \lambda + \ldots$ with coefficients in $K$ such that

$$P(A) = a_0 I + a_1 A + \ldots = 0$$

(1)

We present in this paper a result on these operators (Theorem 1.2). It is established for the case $K = \text{real or complex numbers}$, $E$ a Banach space, $A$ a bounded operator, although it is easily seen to be valid, with an additional assumption, for general $K$, $E$ and $A$. (See the observations after the proof). Theorem 1.2 is proved with the help of a result in [5] (Theorem 15) which, for the sake of completeness, is included here together with Lemma 14 as Theorem 1.1 and Remark 1 respectively. We establish next a version of Theorem 1.2 for certain unbounded operators $A$ (Theorem 2.2) and point out its connections with control theory. Theorem 2.2 is a generalization of Theorem 2.2 of [4] from the case in which the "control" space $F$ has dimension 1 to the case of arbitrary finite dimension.

Paragraph §1 is fairly self contained and makes use only of elementary notions of linear topological algebra; paragraph §2 is

* Most of the results of this paper were obtained at the Courant Institute of Mathematical Sciences, New York University, with the support of a Ford Foundation Pre-Doctoral Fellowship.
closely related to [4], Section § 2 and uses notations, definitions and results in that paper.

§ 1. The case of bounded \( A \)

We shall suppose throughout this paragraph (unless otherwise stated) that \( K \) is the field of real (complex) numbers, \( E \) is a real (complex) Banach space and \( A \) is a bounded operator.

Theorem 1.1. Assume that for every \( u \in E \) there exists a polynomial \( p(\lambda) = p(u; \lambda) \neq 0 \) such that \( p(A)u = 0 \). Then \( A \) is a Jordan operator.

Proof: Let \( p(u; \lambda) \) be the minimal polynomial of \( A \) at \( u \), i.e. the generator of the ideal \( I_u \) (of the ring of polynomials in one indeterminate with coefficients in \( K \)) consisting of all polynomials \( p(\lambda) \) with \( p(A)u = 0 \). Recall that \( p(u; \lambda) \) is uniquely defined, save by multiplication by a nonzero element of \( K \). We have

\[ p(cu; \lambda) = p(u; \lambda) , \ c \in K , \ c \neq 0 \]  
\[ p(u; \lambda) p(v; \lambda) \text{ is divisible by } p(u + v; \lambda) \]  

(1.1) is clear; (1.2) follows from the relation
\[ p(u; A) p(v; A) (u + v) = p(v; A) p(u; A)u + + p(u; A)p(v; A)v = 0 \]

Let us observe next that the degree of \( p(u; \lambda) \) is bounded independently of \( u \). In fact, let
\[ E_N = \{ u \in E \mid \deg p(u; \lambda) \leq N \} \], and let \( \{ u_n \} \) be a sequence in some \( E_N \) convergent to some element \( u \in E \).
Normalize \( p(u_n; \lambda) = a_0 + a_1 \lambda + \ldots \) by, say, the condition \( |a_0| + |a_1| + \ldots = 1 \). By passing, if necessary to a subsequence we can suppose that \( a_{kn} \to a_k \) as \( n \to \infty \); by the normalization condition \( |a_0| + |a_1| + \ldots = 1 \) and therefore \( p(\lambda) = a_0 + + a_1 \lambda + \ldots \neq 0 \). But
\[ p(A)u = \lim p(u_n; A)u_n = 0 \]

hence \( \deg p(u; \lambda) \leq \deg p(\lambda) \) and \( u \in E_N \). This shows that each \( E_N \) is closed. Since \( \bigcup N E_N = E \) the category theorem of Baire im-
plies that some \( E_N \) contains a sphere, say \( \{ u \in E | |u - u_0| \leq \rho \} \).

But if \( v \) is any element of \( E \) \( p(v;\lambda) = p(\rho v/|v|;\lambda) \) divides \( p(u_0 - \rho v/|v|;\lambda) p(u_0;\lambda) \) which shows that

\[
\deg p(v;\lambda) \leq 2N.
\]

Let us pass now to the construction of the polynomial \( P \) in (1). Choose \( u \in E \) such that

\[
\deg p(u;\lambda) = \sup \{ \deg p(v;\lambda) ; v \in E \} \quad (1.3)
\]

We shall show that \( p(u;\lambda) = P(\lambda) \). In fact, let \( w \) be any element of \( E \) such that \( p(w;\lambda) = p_0(\lambda)^m, m \geq 1 \) where \( p_0(\lambda) \) is an irreducible polynomial. In view of (1.2) we have

\[
\begin{align*}
p(u + w;\lambda) p(w;\lambda) & = p(u;\lambda) q(\lambda) \quad (1.4) \\
p(u;\lambda) p(w;\lambda) & = p(u + w;\lambda) r(\lambda) \quad (1.5)
\end{align*}
\]

where \( q, r \), are polynomials. We get from (1.4) and (1.5) that

\[
q(\lambda) r(\lambda) = p(w;\lambda)^2 = p_0(\lambda)^{2m}
\]

so \( q(\lambda) = p_0(\lambda)^k, r(\lambda) = p_0(\lambda)^j, k, j \geq 0, k + j = 2m \)

Then

\[
p(u;\lambda) = p(u + w;\lambda) p_0(\lambda)^h,
\]

\(-m \leq h \leq m. By virtue of (1.3) h \geq 0. But then p(u;\lambda) w = = p_0(A)^h p(u + w;A) (u + w) - p(u;\lambda) u = 0, so p(u;\lambda) is divisible by p(w;\lambda).

Let now \( v \) be any element of \( E \), \( p(v;\lambda) = \Pi_{k-1} p_k(\lambda)^{m_k} \) where \( p_1, \ldots, p_n \) are different irreducible polynomials. It is plain that if \( w = \Pi_{k \neq 1} p_k(\lambda)^{m_k}, p(w;\lambda) = p_1(\lambda)^{m_1} \) By virtue of the preceding considerations \( p(u;\lambda) \) is divisible by all the polynomials \( p_1(\lambda)^{m_1} \), and hence by \( p(v;\lambda) \) itself. This ends the proof of Theorem 1.

Remark 1 Clearly, Theorem 1.1 remains valid for general \( K, E \) and \( A \) if we assume

\[
\sup \{ \deg p(u;\lambda) ; u \in E \} < \infty \quad (1.6)
\]
On the other hand, if (1.6) is false the conclusion of Theorem 1.1 might not hold. In fact, let $E$ consist of all sequences \{ $a_0, a_1, \ldots$ \} of elements of $K$ such that $a_k = 0$ except for a finite number of indices, $A \{ a_0, a_1, \ldots \} = \{ a_1, a_2, \ldots \}$. Then for each $u \in E$ there exists $n = n(u)$ such that $A^n u = 0$; however, it is easy to see that $A$ is not a Jordan operator.

**Remark 2** We only need to assume in Theorem 1.1 the existence of a function $f(u; \lambda)$ for each $u \in E$, analytic in $\sigma(A)$ such that $f(A) u = 0$. In fact, any such $f$ can be written $f = gp$, where $g$ has no zeros in $\sigma(A)$ and $p$ is a polynomial. Then $f(A) = g(A) p(A)$ and, since $g(A)$ is one-to-one $f(A) u = 0$ implies $p(A) u = 0$.

**Remark 3** It is clear from the proof of Theorem 1.1 that we need to assume the existence of $p(u; \lambda)$ (or $f(u; \lambda)$, see Remark 2) only for $u$ in a subspace of the second category of $E$.

**Theorem 1.2** Let $m \geq 1$. Assume that for every $m$-ple $(u_1, u_2, \ldots, u_m)$ there exists a $m$-ple of polynomials $(p_1, \ldots, p_m)$ not all zero such that $\sum_{k=1}^m p_k(A) u_k = 0$. Then $A$ is a Jordan operator.

**Proof:** Let $E^m$ be the Banach space of all $m$-uples $(u_1, u_2, \ldots, u_m)$ of elements of $E$ (pointwise operations) normed with, say, $|(u_1, u_2, \ldots, u_m)| = \max(|u_1|, |u_2|, \ldots, |u_m|)$. Let $E_{E^m}^m = \{ (u_1, u_2, \ldots, u_m) \in E^m$ such that there exists polynomials $p_1, p_2, \ldots, p_m$ not all zero with $\sum_{k=1}^m p_k(A) u_k = 0$ and $\max_{k} \deg p_k \leq N \}$. It is easy to show like in the proof of Theorem 1.1 that each $E_{E^m}^m$ is closed; thus by Baire's category theorem some $E_{E^m}^m$ contains a sphere. This implies again that the degree of the polynomials $p_1, p_2, \ldots, p_m$ in the statement of Theorem 1.2 can be supposed bounded by a constant $N$ independent of $(u_1, u_2, \ldots, u_m)$.

We end now the proof by induction. If $m = 1$ we are in the case considered in Theorem 1.1. Let $m > 1$ and let $(u_1, u_2, \ldots, u_{m-1})$ be any $(m - 1)$-ple of elements of $E$.

Consider the $m$-ple $(u_1, u_2, \ldots, A^{m-1} u_{m-1}, u_{m-1})$.

By the preceding considerations, there exists a $m$-ple $(p_1, p_2, \ldots, p_m)$ of polynomials, not all zero and such that $\sum_{k=1}^{m} p_k(A) u_k + (p_{m-1}(A) A^{m-1} + p_m(A)) u_{m-1} = 0$, $\max_k \deg p_k \leq N$. Since

(*) See [2], VII for the necessary notions of operational calculus.
deg \( p_m \leq N \) the polynomials above cannot be all zero, and thus our inductive step is achieved. Theorem 1.2 is proved.

Remarks 1 and 3 after Theorem 1.1 have evident generalizations to this case. As regards to Remark 2 we only need to assume in Theorem 1.2 for each \((u_1, \ldots, u_m) \in E^m\) the existence of \(m\) functions \(f_1, \ldots, f_m\), analytic in a domain \(D \supset \sigma(A)\) (independent of \((u_1, \ldots, u_m)\)), not all zero, such that \(\mathcal{X} f_k(A)u_k = 0\). The proof is substantially similar to that of Theorem 2.2 below.

§ 2. Rigidity of linear control systems

We consider in this paragraph linear control systems

\[
u'(t) = Au(t) + Bf(t), t \geq 0
\]  

(2.1)

Here \(A\) is the infinitesimal generator of a strongly continuous semigroup \(T(t)\) of bounded operators in the complex Banach space \(E\), \(u(t)\) is a point in the space \(E\) describing the state of the system at the time \(t\), \(f(t)\) is a function (the input or control) with values in some other Banach space \(F\) and the linear bounded operator \(B : F \to E\) is a "transmission mechanism" through which \(f\) acts on (2.1).

We shall understand by a solution of (2.1) with initial data \(u(0) = u_0 E\) and input \(f\) in some space \(L^p(0, \infty ; F)\), \(1 \leq p \leq \infty\), the expression

\[
u(t) = T(t)u + \int_0^t T(t-s)Bf(s)\,ds
\]  

(2.2)

where \(T(t)\) is the semigroup generated by \(A\) (see [4])

A point \(v \in E\) will be called reachable from \(u\) if there exists \(f\) such that the solution \(u(t)\) of (2.1) starting at \(u\) (say, for \(t = 0\)) satisfies \(u(t) = v\) for some \(t \geq 0\).

Definition The linear control system (2.1) will be called rigid if any point \(v\), reachable from another point \(u\) in the time \(t\) by means of some control \(f\) is not reachable from \(u\) in the same time by any control different from \(f\).

It follows easily from the representation (2.2) for the solu-
tion of (2.1) (and the replacement of \( t \rightarrow s \) by \( s \) in the integral) that the system (2.1) will be rigid if and only if the map

\[
f \rightarrow \int_0^t T(s) Bf(s) \, ds
\]

from \( L^p(0, t; F) \) to \( E \) is one-to-one for all \( t > 0 \).

Let us pass now to establish the relation between these notions and the results in § 1. In view of the last observation in the proof of Theorem 2.2 in [4] we need only to consider the case \( p = 2 \).

Observe next that if \( F \) is \( m \)-dimensional unitary space, the space \( \mathcal{L}(F; E) \) of all linear bounded operators from \( F \) to \( E \) can be algebraically and topologically identified with the space \( E^m \) defined in the proof of Theorem 1.2 by means of the correspondence that assigns to the element \((u_1, \ldots, u_m) \in E^m \) the operator in \( \mathcal{L}(F; E) \)

\[
B(x_1, \ldots, x_m) = \sum_{k=1}^m x_k u_k, \quad (x_1, \ldots, x_m) \in F
\]

It is a consequence of the functional calculus for infinitesimal generators (see [4], § 2) that if

\[
f(s) = (f_1(s), \ldots, f_m(s)) \in L^2(0, \infty; F)
\]

and \( B \) is the operator (2.4)

\[
\int_0^t T(s) Bf(s) \, ds = \sum_{k=1}^m \hat{f}_k(A) u_k
\]

where the functions \( \hat{f}_k \) (the Fourier transforms \( \hat{f}_k(\lambda) = \int f_k(s) \exp(\lambda s) \, ds \)) of \( f_k \) belong to the space \( H^2 \) of the left half-plane (see [4], § 2 and [3]).

Finally, let us recall the notion of operator of admissible meromorphic type, generalization of that of Jordan operator for the unbounded case. An infinitesimal generator \( A \) is said to be of admissible meromorphic type if the resolvent \( R(\lambda; A) \) is a meromorphic function with poles of order \( m_k \) at points \( \lambda_k \) and

\[ -\sum m_k \Re \lambda_k / (1 + |\lambda_k|^2) < \infty \]

(see again [4], § 2). The preceding considerations make clear the equivalence of

(*) We endow \( \mathcal{L}(F; E) \) with the uniform topology of operators.
Theorem 2.1. Let $A$ be an infinitesimal generator satisfying conditions (2.1.a), (2.1.b) of [4], § 2. Assume $A$ is not of admissible meromorphic type. Then the linear control system (2.1) is rigid for all operators $B \in \mathcal{L}(F; E)$ except for those in a subset of the first category of $\mathcal{L}(F; E)$ and

Auxiliary Theorem 2.2 Let $A$ satisfy the same conditions of Theorem 2.1. Assume there exists a subset $L$ of the second category of $E^n$ such that for every $(u_1, \ldots, u_n) \in L$ there exist $m$ functions $f_1, \ldots, f_m$ in $H^2$, not all zero and such that $\sum_{k=1}^m f_k(A) u_k = 0$. Then $A$ is of admissible meromorphic type.

For the proof, we shall make use of

Lemma 2.3. Let $\{f_n\}$ be a sequence in $H^2$ of the half-plane $\Re \lambda \leq 0$ such that $|f_n|_m \leq 1$. Then there exists a subsequence $\{f_m\}$ such that:

(a) $\{f_m\}$ converges weakly to a function $f \in H^2$, $|f|_m \leq 1$

(b) $f_m(A)$ converges to $f(A)$ in the uniform topology of operators.

Proof: The fact that there exists a subsequence $\{f_m\}$ satisfying (a) is an elementary fact of the theory of $H^2$ (in fact, Hilbert) spaces. To show (b), let us consider the representation (2.11) of [4]

$$f_m(A) = \frac{1}{2\pi i} \int_{P(c, \theta)} f_m(\lambda) R(\lambda; A) \, d\lambda$$

(2.5)

where $P(c, \theta)$ is the contour $c + |y| \cot \theta + iy, -\infty < y < \infty$ for suitable $c < 0$, $\theta > \pi/2$ (see [4], § 2). Cauchy's formula

$$f_m(\lambda) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{f_m(it)}{it - \lambda} \, dt$$

(2.6)

and the weak convergence of $\{f_m\}$ imply that $\{f_m\}$ converges uniformly on compacts of $\Re \lambda < 0$ to $f$. It is then clear that (b) will hold for $\{f_m\}$ if we can show

$$\lim_{n \to \infty} | \int_{P(c, n, \theta)} f_m(\lambda) R(\lambda; A) \, d\lambda | = 0$$

(2.7)

uniformly with respect to $m$, where $P(c, n, \theta)$ is the intersection
of $P(c, \theta)$ with the region $|\lambda| \geq n$. But in view of (2.1.b) of [4]

$$|R(\lambda; A)| \leq C/|\lambda|$$

for $\lambda \in P(c, \theta)$ and some constant $C$. Furthermore, (2.6) and the Cauchy-Schwarz inequality imply

$$|f_m(\lambda)| \leq C/|\lambda|^{1/2}$$

(2.9) for $\lambda \in P(c, \theta)$ and some constant $C$, uniformly with respect to $m$. (2.8) together with (2.9) imply 2.7) and, a fortiori, Lemma 2.3

Proof of Theorem 2.2. Define subsets $L_{M,N}$ ($M = 1, 2, 3, \ldots, N = 1, 2, \ldots, m$) of $L$ as follows: $L_{M,N} = (u_1, \ldots, u_m) \in L$ such that there exist functions $f_1, f_2, \ldots, f_m$ in $H^2$, not all zero and such that (a) $\max_k |f_k| |n^2| \leq 1$ (b) $|f_x(c)| \geq 1/M$, $c$ a fixed point outside $\sigma(A)$, (c) $\sum_{k=1}^m f_k(A) u_k = 0$. It is easy to see that every $(u_1, \ldots, u_m) \in L$ belongs to some $L_{M,N}$ (if the corresponding functions $f_1, \ldots, f_m$ all vanish at $c$ multiply them by a convenient power of $(\lambda - c)^{-1}$ and that each $L_{M,N}$ is closed (to do this we proceed in a way similar to that of Theorem 1.2 and make use of Lemma 2.3) Again by an application of Baire's theorem we deduce that some $L_{M,N}$ has an interior point, and this can be easily seen to imply (possibly after a rearrangement of indices) that the functions $f_1, \ldots, f_m$ in the statement of Theorem 2.2 can be chosen in such a way that $f_m(c) \neq 0$.

The proof ends now like that of Theorem 1.2. Let $(u_1, \ldots, u_m)$ be any $(m - 1)$ - ple of elements of $E$, and let $g(\lambda) = (\lambda - c)(\lambda + c)^{-2} \in H^2$. Then, if $f_1, \ldots, f_m$ are the functions corresponding to the $m$-ple

$$(u_1, \ldots, u_m, g(A) u_{m-1}, u_{m-1})$$

we have $\sum_{k=1}^{m-2} f_k(A) u_k + (f_{m-2}(A) g(A) + f_m(A)) u_{m-1} = 0$, the functions $f_1, \ldots, f_{m-2}, f_{m-1} g + f_m$ not all zero. This allows us to reduce the case of $m$-ples to the case of $(m - 1)$ ples, and when $m = 1$ Theorem 2.2 reduces to Theorem 2.2 of [4].

Remark Theorem 1.2 states that when $A$ is not of admissible meromorphic type and $F$ is finite-dimensional then (2.1) is rigid for "most" operators $B$ in $L^2(F; E)$. The situation changes when
$F$ is of infinite dimension; for instance, if $E = F$ it is not difficult to see that (2.1) is not rigid when $B$ has a bounded inverse or is not one-to-one.

BIBLIOGRAPHY